

PROBABILITY AND STATISTICS PROGRAMS

for the BBC Micro

by Dr David Green

with R P Knott, PE Lewis and J Roberts

Student's Notes

BULB TEST

Introduction

A manufacturer of light bulbs is testing them twenty at a time in special bulb testers. The bulbs are left alight until they burn out. It is suspected that some bulb testers are defective and cause the bulbs to burn out in a biased or predictable order.

Run the BULB TEST program and write a report on each of the twelve bulb testers. Some defective bulb testers influence the order in which the bulbs burn out. You must find out which bulb testers are defective and which are OK.

What you have to do

- 1 Load the BULB TEST program and run the instructions. This tells you about the problem and describes the following useful facilities:

F - FREEZE SCREEN L - LABEL BULBS

U - UNFREEZE SCREEN S - SPEED UP

(These facilities are used after a bulb tester has been selected. Just press the single letter required.)

- 2 When you have read the instructions, choose option 2 to run the program.
- 3 Choose the Sound Level (1 is recommended)

- 4 Choose the Difficulty Level (1 is recommended)

You can re-run the same experiment if you wish by using Level 1, as this level will always give the same results. Level 3 gives different results every time.

- 5 Choose the Speed Level (0 or 1 are recommended)

With Speed Level 0 the program stops after each bulb burns out.

With Speed Level 1 the program keeps going slowly.

- 6 Choose each of the bulb testers (1 to 12) in turn, and for each do this:

Watch the burn-out sequence and ask yourself these questions:

- a Is the order of burning out surprising?
- b Is there anything predictable?
- c Is there some sort of pattern?

If the answer to any of these questions is 'Yes!', then perhaps the bulb tester is defective.

- 7 Write your report on each bulb tester on the RESULTS Sheet. Remember you can re-run the program with the same bulb tester until you are satisfied.

Name

Date

BULB TEST RESULTS SHEET

Bulb Tester	REPORT ON BURN-OUT PATTERN	OK or Defective
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

Name

Date

CIA SHEET 1

	Heads	Tails	Number of RUNS	Longest RUN
1 HTHTHTHTHTHTHTHTHT				
2 HHHHHHHHHTTTTTTTTTT				
3 HHTHHTHHTHHTHHTHHTH				
4 HTTHHTTTTHTTHTTTTHTT				
5 THHTHTTTHHTHTHHTHTTH				
6 HHHHHTTTTTTHHHTTHHHH				
7				

Comments:

1 _____

2 _____

3 _____

4 _____

5 _____

6 _____

7 _____

CIA SHEET 2

What you have to do

- 1 Get a copy of CIA Sheet 3. Fill in your name and the date.

- 2 EITHER Toss a coin 100 times, recording the results on Sheet 3 Table 1,

OR Run the CIA program:
 - a) Select machine 9 (the 'PLAYFAIR' machine)
 - b) Ask for 100 tosses
 - c) Record the results on Sheet 3 Table 1 (if you wish to keep a record).

- 3 Analyse the data in Table 1 to fill in Table 2.

Use either Method 1 or Method 2 or Method 3 (your teacher will advise you) :

Method 1 If you have used the CIA program you can get it to do the analysis straightaway using the 'Count' and 'Runs' commands.

Method 2 If you have NOT used the CIA program you can still use it to analyse the data, if you wish. Do it like this:

Run the CIA program

- a) Select machine 0 (the 'Auto' machine)
- b) Type in the 100 results
- c) Press E (to End)
- d) Analyse the data using the 'Count' and 'Runs' commands.

Method 3 If you are NOT going to use the CIA program you must do the analysis yourself by counting.

Name

Date

CIA SHEET 3 - DATA SHEET

1-10									
11-20									
21-30									
31-40									
41-50									
51-60									
61-70									
71-80									
81-90									
91-100									

TABLE 1

Number of Heads

Number of Tails

Number of Runs

Longest Run

Number of Runs of Length 1

Number of Runs of Length 2

Number of Runs of Length 3

Number of Runs of Length 4+

TABLE 2

CIA SHEET 4: LIST OF COMMANDS

Key Name	Description of Command
B	Barchart Draws a barchart of the data.
C	Count Counts the number of Heads and Tails on each line and the totals for all the data.
G	Group Groups the data, for example in threes or fives. Groups of size 1 to 10 are allowed.
H	Help Gives descriptions of all the Commands.
L	List Lists the data, 10 tosses per line.
M	Machine Lists the machines ready for one to be selected and used.
P	Pairs Groups the data in pairs and then draws a barchart to show the numbers of HH, HT, TH, and TT.
Q	Quit Finishes running the program.
R	Runs This counts the total number of runs and finds the longest run. Also it counts how many runs are of length 1, and how many are of length 2, etc. <u>Example:</u> HTTTHH has 3 runs: H then TTT then HH. They are of lengths 1 and 3 and 2 respectively.
T	Table Displays table of the data.

NOTES:

- 1) If you press BREAK, the program will be called up again and you will have to start afresh. The data will be lost.
- 2) When you type in a command (e.g., 'B') or select a machine (e.g., '3') there is no need to press RETURN.

When you type in how many tosses (2 to 200) or the group size (1 to 10), then you must press RETURN.

CIA SHEET 5

Judging the Cointossing Machines

The Cointosser Inventors Association have sponsored a Great Cointossing Machine Competition.

You have been appointed to judge these machines. You must detect any weaknesses. A machine that produces Tails every time would not be much good! Nor would a machine that went THTHTHTHTH...!

The tossing must be RANDOM. First then, you need to think about what is meant by RANDOM, and what is not RANDOM.

To help you test the machines you have your microcomputer. The commands to control it are listed on Sheet 4.

What you have to do

- 1 Run the CIA program.
- 2 Select a machine.
- 3 Get it to 'toss' some coins.
- 4 Record the data using CIA Sheet 6, and analyse the results.
- 5 Decide if it is fair or not.
- 6 Write one or two sentences on CIA Sheet 7 - the REPORT Sheet - reporting on anything wrong with the machine.
- 7 Repeat for all eight machines.
- 8 Finally, select the machine or machines which you think are all right and can be used for selecting Premium Bond winners and write their names on the REPORT Sheet.

CIA SHEET 7 - REPORT SHEET

	Machine	Report on Machine
1	Bonker	
2	Dumbo	
3	Flipper	
4	Loopy	
5	Robot	
6	Super	
7	Turbo	
8	Whizz	

Final Judgement Machines suitable for Premium Bond selection

Signed _____

Date _____

THE COSMIC INFORMATION DETECTOR

The Problem

Signals are being received from outer space, in binary digit form, e.g. 0011011010111....

The question is whether or not these strings of binary digits indicate the presence of intelligent beings. If the strings of bits contain information, then life must be present. If the strings are just 'random noise' then there is no evidence for life in outer space.

The Task

You are asked to investigate strings of binary digits coming from several different sources, and to reach a decision about each source.

What you have to do

- a) Run the CID program, analyse the data from each of the sources (A to I) and decide which are non-random sequences and therefore which are indicating the presence of life.
- b) Write up a report, giving reasons for your decisions.

The Tools for Analysing the Data

A set of single-key computer commands is available with the Cosmic Information Detector program:

<u>Key Name</u>	<u>Command Description</u>
S Source	Selects source whose coded data is to be analysed.
B Barchart	Draws barchart of data.
G Group	Groups data in sizes 2 to 5 or performs Pairs or Triples analysis.
L List	Lists data, 10 bits per line.
P Probability	Calculates whether or not probability exceeds 1% or 5%.
Q Quit	Finishes program.
R Runs	Counts number of runs and presents frequency table for run lengths.
T Table	Displays frequency table for data.

DOUBLE-DOT

Load the program Double-Dot and run the instructions, then choose option 2 to play the game:

- 1 a) Choose a grid with 20 dots - for example 4 by 5.
- b) Run the game 100 times.
- c) Make a prediction as to the number of counters the computer will need to finish the game.
- d) Draw a bar chart to show the number of counters needed to finish the game each time. Set it out as shown below:



Look at the table of results and answer these questions:

- e) What is the maximum number of counters needed, in your set of 100 results?
- f) What is the minimum?

- g) What is the average?
- h) How do these results and the bar chart's shape compare with what you expected?

2 A girl is collecting photos of pop stars given away with chewing gum. One photo is given free with every packet. There are 20 photos in the set.

Using your results from Question 1:

- a) Estimate how many packets of gum the girl must expect to buy before she gets two the same.
- b) If she were lucky how many packets might she need to buy to get a complete set?
- c) What about if she were unlucky? Or if the gum producers cheated?

3 A boy is collecting tea cards which are given away, one with each packet of PG Tips tea. There are 25 cards in the complete set. After he has got only six cards he gets one of them again. He thinks that is unlikely. Do you agree? (Provide some evidence to support your answer.)

4 If a room contains just two people then it is unlikely that they will have been born on the same day of the week (e.g., Wednesday). If there are eight people, then it is 100% certain that at least two of them share the same day of the week. How many people are needed to make it more likely than not that there will be a shared day?

- 5 If a room contains just two people then it is unlikely that they were born in the same month (or are of the same sign of the Zodiac). How many people are needed to make a match more likely than not?
- 6 If a room contains just two people then it is very unlikely that they share the same birthday (e.g., 16 July). How many people are needed for a match to be more likely than not?

MULTI-DOT

Load the program Multi-Dot and run the instructions then play the game:

A dot is chosen at random and a counter is placed over the dot. This process is repeated until a predetermined number of counters, n , have been placed over one dot, i.e., the game ends when a pile of counters on any dot reaches a size n . The number n can be chosen to be from 2 to 9. The case $n = 2$ is equivalent to the simpler game "Double-Dot".

What the player has to do is to predict accurately how many counters are required to complete the game. The number needed can, of course, vary; it could be as low as n or as high as n times the number of dots in the grid plus one. Both extremes are unlikely!

Accuracy of prediction

The computer will give your prediction an accuracy score in percentage form, 100% being exactly right.

Getting started

It is wise to begin with a small grid (e.g., 2×2) and work through increasing pile sizes before dealing with larger grids.

Recording results

You should keep a record of your results on the RESULTS sheet. Then you can draw graphs on graph paper to show the relationship between 'grid size', 'pile size' and 'number of counters needed'.

Running the program

The following inputs are requested:

Width of grid	(1 - 20)
Height of grid	(2 - 20)
Number of games	(1 - 100)
Speed	(1 - 5)
Pile size to end game	(2 - 9)
Is screen display wanted?	(This can be switched off to speed up the computations)
Is printer output wanted?	(You won't usually want this)
Prediction?	

Results

Following each run the following results are displayed:

Grid size:
Pile size to finish game:
Games played:
Maximum counters needed:
Minimum counters needed:
Mean number of counters:
Standard deviation:
Your prediction:
Your accuracy:

Frequency Table of results (optional)

Name

Date

MULTI-DOT RESULTS SHEET

	Grid Width	Grid Height	No. of dots in Grid	Pile Size	Games	Results
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

LIGHTS

Description

The computer will display a large number of lights. These are arranged in banks of 6, and there are 30 such banks. This is what might be found in a large factory or office.

Light bulbs burn out after some time in an unpredictable way. It is cheaper to leave the burnt out bulbs and then after some time to replace ALL the bulbs with a new set. This might be done every 6 months, or when 20% of bulbs are burned out, for example.

This computer program simulates the burning out of the light bulbs. You can choose how many bulbs must be burned out before replacement should take place. (This is called the replacement level.)

Stage 1

- a) Get three copies of the LIGHTS DATA SHEET.
- b) Run the LIGHTS program. Before you choose a replacement level, answer this question:

"If 30 light bulbs are to burn out, how many of the 30 banks of 6 will have:

- none burned out?
- just 1 burned out?
- 2 or more burned out?

Enter your prediction in TABLE 1:

	Banks with None burnt out	Banks with 1 burnt out	Banks with 2 burnt out	Banks with 3 or more burnt out
Prediction				
Actual results				
<u>Comment</u>			

TABLE 1

- c) Now select Run Type 2, so that you get different results each time, and select replacement level 30.
- d) Make a note of the actual results, in Table 1.
- e) Were your predictions accurate?
Are you surprised at what happened?
Write brief comments in Table 1.

Stage 2

- f) Given your experience from Stage 1 now repeat the experiment once more. Put your predictions, the actual results, and your comments in Table 2.

	Banks with None burnt out	Banks with 1 burnt out	Banks with 2 burnt out	Banks with 3 or more burnt out
Prediction				
Actual results				
<u>Comment</u>				
.....				
.....				

TABLE 2: Replacement level 30

- g) Now run the program 10 times, with replacement level 30 selected each time. Enter all your results in the table on a DATA SHEET, giving you 10 sets of data in all. (Call it Table A.)
- h) Add up the columns to complete Table A.
- i) Plot a bar chart on the same DATA SHEET to illustrate these results and label it Graph A.
- j) Using the results in Table A, answer the following seven questions, putting your answers in Table 3.

What number and what percentage of the banks of lights have:

- 1) No dead bulbs?
- 2) 1 dead bulb?
- 3) 2 dead bulbs?
- 4) 3 dead bulbs?
- 5) 4 dead bulbs?
- 6) 5 dead bulbs?
- 7) 6 dead bulbs?

<u>Number of lights burnt out: 30</u>							
Number of dead bulbs:	0	1	2	3	4	5	6
Number of banks:
Percentage of banks:

TABLE 3

Stage 3

- k) Now predict the results if a replacement level of 60 is selected. Put your predictions in Table 4.

	Banks with None burnt out	Banks with 1 burnt out	Banks with 2 burnt out	Banks with 3 or more burnt out
Prediction				
Actual results				
<u>Comment</u>			

TABLE 4: Replacement level 60

- l) Now perform the experiment, recording your results on a DATA SHEET, labelling the entries Table B and the bar chart Graph B.
- m) Fill in the 'Actual results' section of Table 4.
- n) Fill in and complete Table 5, using the results from Table B.

<u>Number of lights burnt out: 60</u>							
Number of dead bulbs:	0	1	2	3	4	5	6
Number of banks:
Percentage of banks:

TABLE 5

Stage 4

- o) Now predict the results if a replacement level of 120 is selected. Put your predictions in Table 6.

	Banks with None burnt out	Banks with 1 burnt out	Banks with 2 burnt out	Banks with 3 or more burnt out
Prediction				
Actual results				
Comment			

TABLE 6: Replacement level 120

- p) Now perform the experiment, recording your results on a DATA SHEET, producing Table C and Graph C.
- q) Complete Table 6.
- r) Using Table C, complete Table 7.

Number of lights burnt out: 120

	0	1	2	3	4	5	6
Number of dead bulbs:
Number of banks:
Percentage of banks:

TABLE 7

Stage 5

- s) Compare the three sets of results, for replacement levels 30, 60 and 120 as set out in Tables 3, 5 and 7. Write a short report on what you notice:

Comparison of the 3 sets of results

- t) If each bank of lights were situated in a different office in a large firm of accountants (for example), how many bulbs should be allowed to burn out before all bulbs in the firm are replaced? Write a short report explaining your answer:

Report on a suitable replacement level

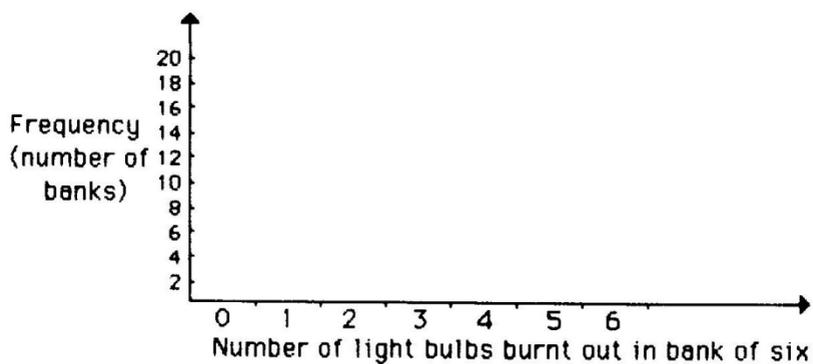
Name

Date

LIGHTS DATA SHEET

Experiment	Number of banks which have these numbers of light bulbs burnt out						
	0	1	2	3	4	5	6
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
TOTAL							

TABLE



GRAPH

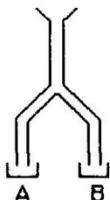
ROLLERBALL

Load the program ROLLERBALL and choose option 1 - 'Instructions', then work through the following investigations:

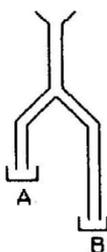
- Using option 3, 'Call Design Program' design each of the following shapes in turn and use the designs to roll 100 balls down each.

Choose the 'Random' option and then the 'Fair' option so that the balls are just as likely to go Left as Right at each junction. Record your results on the Rollerball Results Sheet.

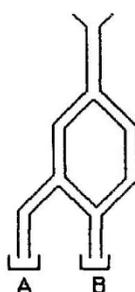
(A)



(B)

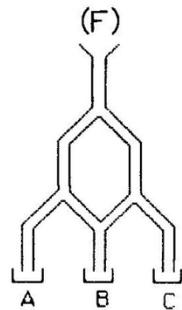
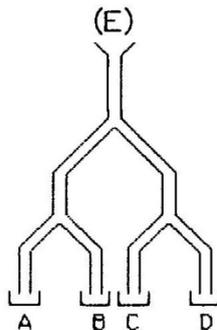
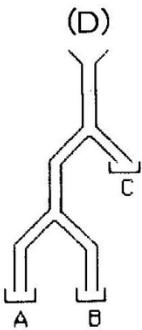


(C)



- Which shapes (if any) gave about equal numbers of balls in boxes A and B?
 - Can you explain why?

- b) Which shapes (if any) gave many more balls in one box than the other? Can you explain why?
- 3 a) Try to predict what will happen when 100 balls are rolled down each of the designs below, when the 'Fair' option is selected.



- b) Design each of the shapes in Question 3a and roll 100 balls down each. Compare the results with your prediction.
- c) Write comments about each set of results and your predictions.
- 4 Repeat Question 3 when the 'Biased' option is selected. First choose the bias to be 25% going Left, then choose the bias to be 75% going Left.
- 5 Repeat Question 4 selecting various other biases. Be sure to write down comments about your predictions and the actual results.
- 6 Make up your own design. Predict how many balls will end up in each box with 100 balls rolled, when the 'Fair' option is chosen, and say why. Then check your prediction and comment on it.
- 7 Repeat Question 6 choosing the 'Biased' option for various biases.

SAMPLER

The computer allows you to:

- 1a choose one of several fixed sets of 120 numbers. This is your 'parent population'. It displays that set of numbers and plots its histogram (in yellow). This graph shows the parent distribution;
- b decide how many random samples you want to pick from the parent population. You can pick from 1 to 9999 samples, but choosing over 100 samples is not usually a good idea! and
- c decide how large each sample should be, i.e., how many numbers in each sample. You can have from 1 to 100 numbers per sample.
- 2 It calculates the mean and standard deviation of each sample and plots (in red) a histogram of all these sample means. This graph shows the sample mean distribution.
- 3 You can then compare the parent distribution and the sample mean distribution.
- 4 The purpose of this work is to investigate how the sample means are distributed. What shape do we get when we plot the sample means? How does the sample mean distribution compare with the parent distribution.
- 5 Before you start you should think about the following questions, then, to help you answer them fully, you will need to carry out the experiments described in Section 6.

- a What influences the sample mean distribution?
- b Is the sample size important?
- c Is the number of samples taken important?
- d Is the shape (i.e., distribution) of the parent population important?

6 For each of the following experiments:

- a Make a sketch of the parent distribution.
- b Make a note of the position of the centre of the parent distribution and of the sample mean distribution. These are reported on the screen as 'Pop mean' and 'Sample mean'.
- c Make a note of the spread of the parent distribution and of the sample mean distribution. These are reported on the screen as 'Pop S.D' and 'Sample S.D'.
- d Make a sketch of the sample mean distribution.

Experiment 1

- a With sample size 1, take 500 samples from the RECTANGULAR distribution.

Compare parent and sample mean distributions.

- b With sample size 10, take 50 samples from the RECTANGULAR distribution.

Compare parent and sample mean distributions.

- c With sample size 50, take 20 samples from the RECTANGULAR distribution.

Compare parent and sample mean distributions.

Experiment 2

Repeat Experiment 1, for the SKEW distribution.

Experiment 3

Repeat Experiment 1, for the BIMODAL distribution.

Experiment 4

Repeat Experiment 1, for the BELL-SHAPED distribution.

Experiment 5

Repeat Experiment 1, for the U-SHAPED distribution.

Experiment 6

Make up your own distribution using the SAMPLE GENERATOR and repeat Experiment 1 for that.

7 The Final Report

Write a report, using the information gained from the experiments, to answer the series of questions set out in Section 5.

SEEDS

Description

The computer will show you 12 pictures of seeds which have grown into flowers. Many thousands of seeds have been planted in the seed boxes but only a few have grown.

The pictures are supposed to be realistic, but are they? You should find something unusual about most of the pictures.

Stage 1

- a Run the SEEDS program and select 'SEEDS A Instructions'. Run this through until you are ready to proceed.
- b Run 'SEEDS A', and for each of the 12 pictures answer the three questions below.

Question 1

Are the positions of the flowers suspicious?

- Choose from:
- a) Yes
 - b) No
 - c) Unsure

Question 2

How surprised would you be to find the flowers grow up in the pattern shown?

- Choose from:
- a) Very surprised
 - b) Quite surprised
 - c) A bit surprised
 - d) Not at all surprised
 - e) Unsure

Question 3

What is surprising about the pattern produced?

Answer in your own words.

Write your answers in Table A on the SEEDS result sheet.

Stage 2

- a Run the SEEDS program and select 'SEEDS B Instructions'.
- b When you are ready select 'SEEDS B'.

You should run the program through at least three times for each picture. You will find the pattern may be different each time for the same picture.

For each of the 12 pictures answer the three questions in Stage 1.

Write your answers in Table B on the SEEDS results sheet.

Name

Date

SEEDS RESULTS SHEET -TABLE A

Picture	Suspicious?	How surprised?	Reasons
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

Name

Date

SEEDS RESULTS SHEET -TABLE B

Picture	Suspicious?	How surprised?	Reasons
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

